

Enfoques en torno al modelo cognitivo para la recuperación de información: análisis crítico

Benjamín Vargas-Quesada

E-mail: bvargas@ugr.es

Félix de Moya Anegón

E-mail: felix@ugr.es

M^a Dolores Olvera Lobo

E-mail: molvera@ugr.es

Departamento de Biblioteconomía y Documentación
Facultad de Biblioteconomía y Documentación
Universidad de Granada.

INTRODUCCIÓN

Vivimos en una sociedad que se caracteriza por la inmensa cantidad de información que produce y por la accesibilidad de esta información para toda persona que disponga de los medios necesarios y esté interesada en ella. Esto ocurre, en gran medida, gracias al desarrollo de las nuevas tecnologías y, con ellas, a la amplia proyección de Internet, de las bibliotecas digitales y de todo tipo de bases de datos en línea. Desdichadamente, los distintos modelos de recuperación de información existentes, y los diversos sistemas de recuperación de información a los que éstos han dado lugar, no han mejorado ni evolucionado en la medida que demandaba el crecimiento de la información y del acceso a la misma.

Consideramos *modelo* a aquel esquema teórico de un sistema o de una realidad compleja, que se elabora para facilitar su comprensión y el estudio de su comportamiento. A finales de los años 50 tiene lugar un hito de indudable repercusión en el ámbito de la recuperación de información (en adelante, RI). El proyecto Cranfield desarrollado por Cyril Cleverdon[1] en dos etapas (Cranfield I y Cranfield II) es, en sentido estricto, la primera iniciativa dirigida a crear un modelo – conocido como modelo o aproximación tradicional o algorítmica –, una metodología y unas pautas para la evaluación de los sistemas de recuperación de información todavía vigentes.

De esta manera, la recuperación de la información, se concibe como una actividad estática, de procesos “batch” o por lotes[2],[3] en los que el usuario realiza una consulta y el sistema responde con un listado de documentos. Este modelo parte de que la información procesada por el sistema informático se organiza en forma de documentos, que constituyen las representaciones documentales, entendiendo éstas como un conjunto de caracteres que se agrupan formando palabras, que a su vez se agrupan formando frases y por último párrafos y que éstos de forma más o menos extensa componen los documentos[4]. Por otro lado está la consulta del usuario, que se compone de uno o más términos y que representa su necesidad informativa. Básicamente el proceso es muy sencillo, se trata de equiparar la representación de los documentos almacenados en la base de datos del sistema de

Resumen

El modelo de recuperación de información tradicional o algorítmico ha ido evolucionando hasta dar lugar al modelo cognitivo. A partir de éste último han surgido distintas propuestas de implementación práctica siendo las más destacables las denominadas por sus autores Modelo Global de Polirepresentación, Modelo Episódico, Modelo Estratificado y Modelo de Retroalimentación Interactiva, las cuales en este trabajo se consideran sistemas y se describen y analizan desde un punto de vista crítico.

Palabras-clave

Recuperación de información; Modelo cognitivo.

Perspectives around the cognitive model for the information retrieval: a critical analysis

Abstract

The traditional or algorithmic model of information retrieval has evolved into the cognitive model. From the latter, different proposals for practical implementation have arisen being the most remarkable ones those labelled by their own authors as Global Model of Polirepresentation, Episodic Model, Stratified Model and Model of Interactive Feedback. These are critically described and analysed in this paper

Keywords

Information retrieval; The cognitive model.

recuperación de información (en adelante, SRI) con la representación de la necesidad informativa del usuario, esto es, la consulta. El problema se presenta cuando no todo lo recuperado por el SRI es adecuado para las consultas realizadas por los sujetos[5].

El diseño de SRI experimentales como el SMART de Salton[6; 7], el desarrollo de técnicas de retroalimentación por relevancia (*relevance feedback*) como las propuestas por Rocchio[8], o los análisis del concepto de relevancia como los realizados por Saracevic[9] y Mizzaro[10], no son más que una pequeña muestra de que el modelo utilizado y sus técnicas derivadas, no eran perfectas.

A mediados de los 70, y como consecuencia del desarrollo de la informática y de las tecnologías de la comunicación, la RI comienza a convertirse en un proceso interactivo. Los problemas anteriormente mencionados y la falta de adecuación por parte del modelo tradicional a las nuevas necesidades, hacen que se levanten las primeras voces en la comunidad científica proponiendo un enfoque cognitivo de la documentación[11]. Se considera que la RI es un proceso interactivo y dinámico y que el usuario forma parte integrante de él. Como eco de estas voces, ven la luz los trabajos de Belkin, Oddy y Brooks [12] [13], Bates[14], o Ellis[15], por citar algunos. A pesar de que estas aportaciones abordan el problema desde diferentes puntos de vista y plantean propuestas claramente distintas, todas tienen algo en común que será lo que las distinga del resto de investigaciones realizadas hasta ese momento, esto es, el estudio del usuario desde sus distintas perspectivas.

A finales de los años 80 y primeros de los 90, son numerosos los autores que reclaman más investigaciones basadas en SRI en entornos reales, y no en entornos simulados como se venía haciendo hasta ese momento[16]. Esto incluye al usuario como un elemento más del proceso de RI. Sin embargo, pese a que hoy han transcurrido más de veinticinco años desde la aparición del modelo cognitivo, y aunque se ha recorrido un largo camino desde entonces, éste no parece suponer una auténtica alternativa al modelo algorítmico o tradicional, ni a los problemas a los que se enfrentaba. De hecho, se hace evidente que la RI en entornos reales, en la mayoría de los casos, sigue excluyendo al usuario. A título de ejemplo, la Conferencia de Recuperación de textos TREC (*Text Retrieval Conference*), que parece haber sido, desde varios puntos de vista, reacia a la RI interactiva[17], hasta su quinta edición no la incorpora, como una de sus áreas de investigación, la *TREC Interactive Track*[18].

Para determinar si, ciertamente, el modelo cognitivo supone una auténtica alternativa al modelo tradicional y a los actuales problemas de la RI, se han de considerar las distintas propuestas de implementación realizadas por sus principales representantes a partir del modelo teórico general. Si bien estas propuestas son parciales y no integran todos los aspectos teóricos del modelo cognitivo, es cierto que permiten una primera aproximación práctica al mismo.

MODELO COGNITIVO

El modelo cognitivo engloba lo que constituye la base del modelo tradicional, dando lugar a lo que Ingwersen denomina la *Teoría Cognitiva de la Recuperación de la Información(...)*, que implica un intento de globalizar la recuperación de información, por medio de la representación de todos sus componentes, como representaciones estructurales con distintos niveles de complejidad que cooperan en un proceso de comunicación interactiva[19].

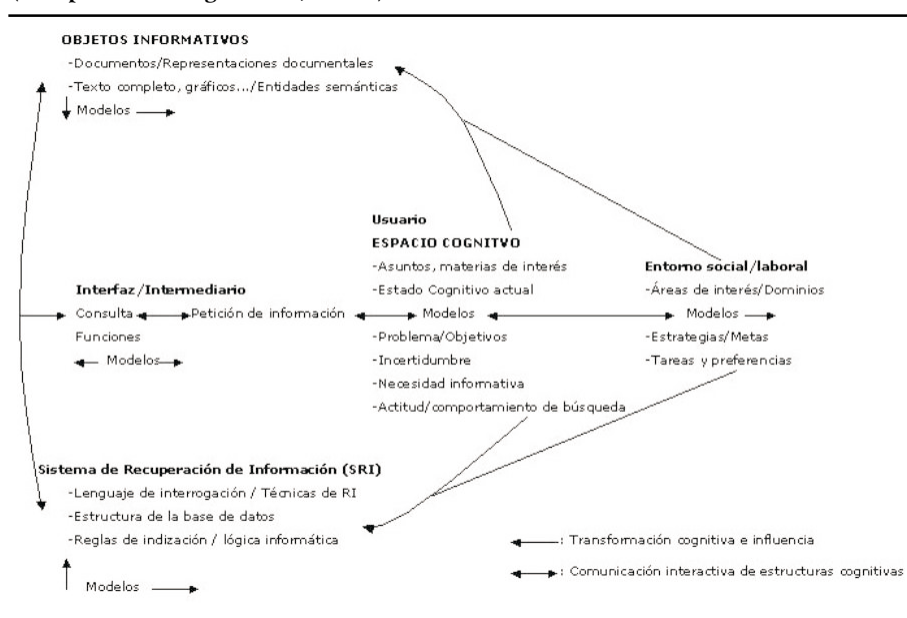
El modelo cognitivo, a diferencia del tradicional, no se pone en marcha a partir del momento en que el usuario realiza una consulta al SRI, sino que comienza incluso antes de que se produzca la necesidad informativa del usuario. En este modelo se tiene en cuenta la extensa gama de factores implicados en el proceso de RI. Son especialmente relevantes aquellos aspectos que influyen en el estado cognitivo y contextual del individuo, por ejemplo, su estado anímico, sus creencias, los objetivos planteados, su grado de motivación o sus áreas de interés, entre otros. Asimismo, interesa estudiar cómo éstos interaccionan entre sí, provocando la necesidad informativa, o lo que Kuhlthau denominó el principio de incertidumbre[20], y Belkin, Oddy y Brooks llamaron ASK[12; 13] (*Anomalous State of Knowledge*) o estado anómalo del conocimiento. Se llama la atención acerca de las dificultades con las que se enfrenta el usuario para conceptualizar y expresar de forma escrita su necesidad informativa, el llamado por Ingwersen[21] *label effect*, es decir, el proceso por el que el usuario *etiqueta* o pone nombre a su necesidad informativa. Esto provoca que cualquier tipo de intencionalidad subyacente que el usuario tenga asociada a su necesidad informativa, se pierda en el momento en el que éste expresa de forma escrita su consulta. Además, la necesidad informativa puede ser variable como consecuencia de la interacción con el SRI[22]. Y por último, aunque no menos importante, en el marco del modelo cognitivo, también éste analiza las técnicas y procedimientos que se aplican en el proceso de RI.

El problema de la recuperación de información interactiva – que es como se conoce al proceso de RI en el modelo cognitivo – ha sido tratado desde diferentes perspectivas a lo largo de su corta historia[23]. Sin embargo, algunos intentos son demasiado específicos, centrándose en aspectos muy concretos del usuario como la formación de la necesidad informativa[12; 13; 20], su naturaleza cambiante[14] o el comportamiento del usuario durante el proceso de la RI[24]. Otros, en su afán por cubrir e integrar todos los aspectos y elementos que intervienen en el proceso de RI[23], presentan un marco teórico tan amplio que en ocasiones resulta bastante difícil poder imaginar su aplicación práctica, y mucho menos su evaluación.

Debido a todos los factores que intervienen en la recuperación de información interactiva (en adelante RII), ésta es tan cambiante que aún no se alcanzan a comprender muchos de los aspectos del proceso de interacción[3]. Éste es el principal motivo por el que la RII, está hoy día en su infancia[5; 25] y con ella el modelo cognitivo de RI. Por ello y, mientras que el modelo tradicional – cuyo proyecto experimental más representativo es el SMART (*System for Manipulation and Retrieval of Text*) y ahora, la Conferencia TREC – durante décadas no ha dejado de evolucionar – sirva como ejemplo el último trabajo de Salton y Buckley[26] –, el modelo interactivo de recuperación de información aún no ha realizado ese recorrido.

Sin duda alguna, es Peter Ingwersen[19; 23; 27], el principal exponente y defensor del modelo cognitivo de RII – también conocido por su nombre, es decir, *modelo de Ingwersen* –. Lo que este autor hace es aglutinar en un único modelo de RI todos aquellos elementos, procesos y técnicas que puedan intervenir en la RII, independientemente de su procedencia. Mediante esta aproximación se da cabida a los actores que intervienen en la RI, a los procesos cognitivos que tienen lugar en el usuario como consecuencia de la aparición de la necesidad informativa y su evolución hasta que ésta se satisface, así como las técnicas y algoritmos necesarios para la RI, entre otros.

FIGURA 1
Modelo Cognitivo de Recuperación Interactiva de Información
(Adaptado de Ingwersen, 1996)



Tal y como se observa en la figura 1, son cinco los elementos principales del proceso de RII, a saber, el usuario, el entorno, la interfaz, el SRI y los objetos informativos.

En el centro, como no podía ser menos, se sitúa el usuario junto con todos aquellos elementos que intervienen en la RII y que la conforman desde un punto de vista cognitivo. Por ejemplo, como se ha comentado anteriormente, el modelo tradicional entiende que necesidad informativa y consulta son exactamente lo mismo. No contempla la posibilidad de que aspectos importantes unidos a la petición de información – y que el individuo no ha sido capaz de conceptualizar o etiquetar –, puedan estar asociados a ella. Como se puede percibir, este modelo no sólo tiene en cuenta al usuario, sino que de éste incorpora su estado cognitivo, contextual y anímico, su necesidad informativa y el nivel de incertidumbre provocado por el ASK, y además analiza cómo interactúa el sujeto con el resto de los componentes del sistema para satisfacer esa necesidad de información.

El entorno social y/o laboral en el que se encuentra inmerso el usuario influye, a veces de forma determinante, en el área de interés del usuario, en sus objetivos o sus preferencias, pudiendo llegar a ser el desencadenante del ASK como se verá más adelante.

En la parte superior izquierda de la figura 1 se representan los objetos informativos propiamente dichos de la base de datos, tales como texto, datos numéricos, ilustraciones, fotografías, sonido, vídeo, diapositivas, metadatos o propiedades asociadas a ellos.

El SRI está compuesto, entre otros elementos, por el lenguaje de interrogación, las técnicas de recuperación de información – integrando de esta forma al modelo tradicional –, la estructura de la base de datos – y a los diseñadores e informáticos que la han realizado – y las reglas de indización – también a los indizadores de los objetos informativos en el caso de que la indización sea manual –.

Por último, las funciones y actores fundamentales de la interfaz o intermediario son la petición de la información y su representación, es decir, la consulta.

En este modelo de RII, la interacción entre los elementos principales tiene lugar cuando confluyen los intereses de cada uno de los actores que intervienen. Por ejemplo, cuando el área de interés o dominio del entorno social/laboral coincide con los asuntos/materias de interés, o problemas/objetivos del espacio cognitivo del usuario, se puede estar ante dos casos: a) Una interacción entre los dos elementos principales del modelo, que inducen a una alteración del estado de conocimiento del usuario. Éstos provocan un estado anómalo de conocimiento, y como consecuencia, una incertidumbre que se manifiesta en forma de necesidad informativa, la cual es transformada en una petición de información, y ésta en una consulta que interactúa con el SRI y con los objetos informativos. b) Una interacción entre los dos elementos principales del modelo, que es rechazado por el espacio cognitivo del usuario y que, al no despertar en él un especial interés o estado anómalo de conocimiento, no da lugar a la necesidad informativa.

El modelo cognitivo de RII es el conglomerado más evolucionado de entre los existentes hasta el momento. Por ejemplo el modelo tradicional está incluido en él por medio de los objetos informativos y del SRI. Además, en la representación esquemática de la figura, se pone claramente de manifiesto una de las principales críticas que se le suelen hacer al modelo tradicional, a saber, que no refleja bien la interacción en el proceso de RII, pues en él sólo intervienen dos de los cinco elementos que lo componen. Por otro lado, en el modelo tradicional, las estructuras documentales y cognitivas que se almacenan en el SRI interactúan unas con otras, o incluso con el usuario, a través de la retroalimentación por relevancia,

pero sólo a un nivel lingüístico o algorítmico, nunca a un nivel cognitivo. En el modelo propuesto por *Ingwersen*, estas interacciones algorítmicas se producen igualmente, pero es el usuario el que, por medio de sus distintos niveles cognitivos y contextuales y de la interacción con los diferentes niveles del sistema, hace que la información obtenida se convierta en nuevos conocimientos.

Las ventajas principales de este modelo se refieren a que muestra de una forma general y sencilla qué es, cómo funciona y qué elementos componen la RII, todo ello desde una sólida base conceptual y en algunos casos empírica.

Por el contrario, sólo cuenta, en nuestra opinión, con un inconveniente, pero que por su importancia resulta crucial. Esto es, el proceso de retroalimentación se contempla como algo implícito al modelo y normalmente asociado a la interacción. En este sentido, debería estar más elaborado y depurado, e incluso habría de contemplar la posibilidad de incluir la retroalimentación como el sexto elemento del proceso de RII.

A partir del modelo general propuesto por *Ingwersen* otros autores han desarrollado diferentes implementaciones prácticas que han denominado *modelos*. No obstante, desde el punto de vista de la teoría de RI y de su terminología, se trata realmente de diferentes sistemas, si bien éstos no han llegado a ver la luz en entornos reales al alcance de los usuarios. Las propuestas más relevantes son las que se recogen a continuación.

MODELO GLOBAL DE POLIREPRESENTACIÓN

La utilización de representaciones múltiples de las consultas y de los documentos así como la aplicación de diferentes técnicas en el proceso de RI, es algo por lo que los investigadores vienen interesándose desde hace tiempo. Esto se debe a un doble motivo. Por un lado la RI es tan compleja e incierta que cualquier representación o técnica de recuperación está diseñada para ofrecer lo mejor de sí misma pero, aún así, sólo abarca una parte de la complejidad informativa, lo que hace necesario la utilización de múltiples técnicas de recuperación y representación. La otra causa tiene su base en el modelo probabilístico de RI de *Robertson*[28] que sugiere que cuantas más pruebas o evidencias tengamos sobre la consulta, sobre los documentos y sobre las relaciones entre ellos, mayores serán las probabilidades de que los resultados se adecúen a la necesidad informativa del usuario[29].

El concepto de polirepresentación es utilizado por Ingwersen[21] para hacer referencia a lo que él denomina redundancia intencional – representación múltiple de las consultas, de los documentos y de las técnicas de RI –, para distinguirlo de la no siempre deseable redundancia general – inconsistencia en la indización y en las técnicas de recuperación –. Son numerosos los estudios realizados, recogidos en el trabajo anteriormente citado, que demuestran que la utilización simultánea de las técnicas de equiparación exacta y parcial producen resultados iguales o mejores que los obtenidos utilizando una sola técnica.

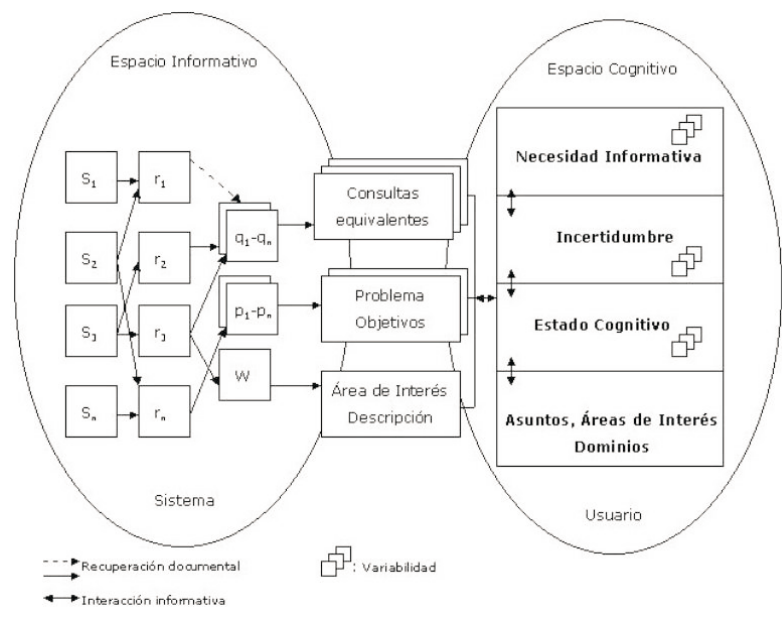
La idea de la polirepresentación, enfocada desde un punto de vista cognitivo, es la que lleva a Peter Ingwersen a afirmar que la inconsistencia de los usuarios al realizar las consultas, la inconsistencia de los indizadores al representar mediante términos el contenido de los documentos y la inconsistencia de las distintas técnicas de RI, puede verse aumentada además por la inconsistencia del usuario para expresar su necesidad informativa. Al igual que en los casos anteriores, se ha comprobado que distintas representaciones de una misma necesidad informativa sometidas a *solapamiento* deben, al menos, ofrecer un resultado igual o mejor que el de la mejor expresión de la necesidad informativa.

Basándose en la idea de la redundancia intencional y en las redes de inferencia de Turtle y Croft[30], Ingwersen propone un sistema de polirepresentación para la RI como apoyo o complemento práctico del modelo cognitivo de RII.

Tal y como se observa en la figura 2, en la derecha, y dentro de lo que se considera el espacio cognitivo del usuario, aparecen ordenados por causalidad, de abajo a arriba, los elementos cognitivos responsables de la formación de la necesidad informativa. Por una parte, hay que destacar que los distintos niveles o elementos interactúan entre sí dando lugar a la aparición o al rechazo de la necesidad informativa y, por otro lado, que el estado de conocimiento, la incertidumbre y la necesidad informativa, por este orden, pueden variar como consecuencia del propio estado cognitivo del usuario y de la interacción con el resto de niveles o elementos cognitivos.

FIGURA 2

Modelo General de Polirepresentación de Recuperación de Información Interactiva (Adaptado de Ingwersen, 1996)



En el centro se encuentra la interfaz o el RMB (*Request Model Builder*) [27]. El RMB es un dispositivo intermediario destinado a detectar la necesidad informativa del usuario, a ayudarlo a conceptualizarla – teniendo en cuenta las variables afectivas, sociales y contextuales que influyen en ella –, y a facilitarle el proceso de *etiquetado* o transcripción de la consulta en la interfaz. Dependiendo de cuál sea el estado de la necesidad informativa detectado por el RMB, llevará a cabo uno de los siguientes procesos: a) realizará una equiparación exacta, parcial, o de ambos tipos mediante distintas técnicas y algoritmos de RI para provocar redundancia intencionada; b) mostrará al usuario necesidades informativas de otros usuarios, muy similares a la suya, para ver si le ayudan a conceptualizarla y etiquetarla, o bien, en el caso de que ésta no aparezca nada clara, c) le propondrá que navegue por la estructura de la base de datos mediante *browsing*, *clustering*, u otro método, pues está demostrado empíricamente que la necesidad informativa, al menos en su fase inicial es a menudo exploratoria, dando lugar a lo que Bates[14] llama “*berry-picking*”.

En el modelo tradicional se acostumbra a preguntarle al usuario por su necesidad informativa, y éste está también habituado a transmitir lo poco que pueda saber de ella.

Sin embargo, el usuario no puede expresar ni transcribir lo que no sabe, lo cual es, realmente, lo que constituye su necesidad informativa. Como mucho, podrá decir qué es lo que no sabe de lo que sabe. Es aquí donde empieza a funcionar el RMB.

La parte izquierda de la figura 2, constituye el espacio informativo del sistema. En él aparecen las representaciones documentales ($S_1, S_2... S_n$), los descriptores o características contenidos en las representaciones de los documentos ($r_1, r_2... r_n$) y, por último, las respuestas obtenidas como consecuencia de aplicar distintas técnicas de recuperación o de representación a la consulta, es decir, polirepresentación o redundancia intencionada ($q_1-q_n, (p_1-p_n), (W)$).

Como se puede observar, la redundancia intencional es fácilmente aplicable en el RMB, bien por medio de la polirepresentación de necesidades informativas similares almacenadas en el sistema y procedentes de consultas realizadas por otros usuarios, bien en el espacio informativo del sistema mediante la aplicación de distintas técnicas de representación documental y de RI.

La ventaja principal de esta propuesta es, al igual que en el modelo cognitivo del cual procede, la facilidad para mostrar la interacción entre los distintos actores de la RII. Además, integra a la polirepresentación o redundancia intencional tanto en la necesidad informativa como en las representaciones documentales y en las técnicas de RI, mediante la fusión de datos.

Su inconveniente es el mismo que se ha señalado para el modelo del que proviene, es decir, la retroalimentación se entiende como algo implícito a la interacción y no está lo suficientemente explicada ni desarrollada. Además, no se detalla la manera en que se producen las distintas iteraciones entre el usuario y el sistema hasta que se consigue la respuesta que más se ajusta a la consulta del usuario. Parece ser que ésto se deja en manos de la retroalimentación por relevancia de las redes de inferencia y en las probabilidades asignadas a sus términos y enlaces.

Trabajos como el de Lee[31], que utiliza con muy buenos resultados la polirepresentación o redundancia intencional en las representaciones documentales y en la RI, o como el de Ruthven, Lamas y Van Rijsbergen[32; 33] que imita la conducta del usuario en el proceso de evaluación de la información mediante el uso de inferencia abductiva, son una breve muestra de que la base del sistema propuesto por Ingwersen es algo más que un principio teórico.

MODELO EPISÓDICO

Belkin denomina *episodios*[34] al conjunto de interacciones que se producen entre el usuario y el sistema durante cada consulta – de forma similar a las que antes tenían lugar entre los usuarios y los intermediarios de las bases de datos – para conceptualizar, etiquetar y transcribir la necesidad informativa, emitir juicios de relevancia sobre uno o varios documentos, u otras.

Al igual que ocurría con el modelo cognitivo, la propuesta de *Belkin* también adopta el apellido de su autor y principal garante. Ésta es la causa de que se haga referencia a él como modelo episódico o de *Belkin*.

Como se observa en la figura 3, en la página siguiente, los elementos o actores que componen este sistema son los mismos que se utilizan en el modelo tradicional o en otros sistemas interactivos, es decir, técnicas de navegación, de visualización, de indización, de representación y de equiparación.

Lo peculiar del sistema basado en episodios es que presta muy poca atención a la estructura de los documentos y a su recuperación, sólo la que le dedica el modelo tradicional, pues se basa en él para realizar este tipo de tareas. Su centro de atención se encuentra en el ASK del individuo, en cómo representarlo y en cómo recuperarlo, pues como se verá más adelante este sistema se basa en el almacenamiento, la recuperación y la interacción de estrategias informativas de búsqueda.

La premisa en la que se basa es que el ASK del individuo es inherentemente difícil de representar, al igual que la consulta, por lo que esta última siempre será una representación aproximada e imperfecta[35]. Por ello, es imprescindible diseñar e implementar una interfaz que sea capaz de detectar la necesidad informativa del usuario e interactuar con su ASK, proponiéndole distintas estrategias que le permitan satisfacer de una forma adecuada su necesidad.

Pueden identificarse dos escuelas o tendencias en relación con el diseño de interfaces interactivas. Una es la que imita el comportamiento de los bibliotecarios o antiguos intermediarios de bases de datos en la búsqueda de información. La otra es la que anima al usuario a interactuar con el sistema mediante un proceso de retroalimentación. La inteligencia artificial está centrando sus esfuerzos en encontrar un equilibrio entre ambas[27]. Lo mismo ha pretendido *Belkin* con el sistema episódico propuesto.

La interacción se realiza por medio de una interfaz *adaptativa*, bastante parecida al RMB de *Ingwersen*. Ésta interactúa con el usuario para ayudarlo a detectar y resolver su necesidad informativa mediante el uso de *scripts* y de reglas conversacionales. Los denominados *scripts*, son la combinación de estrategias de búsqueda que el SRI tiene almacenadas en su base de datos, bien como consecuencia de la interacción con otros usuarios, bien porque el diseñador del sistema las ha implantado antes de su puesta en funcionamiento. La aplicación de reglas conversacionales permiten transformar las estrategias de búsqueda en diálogos estructurados entendibles tanto por el usuario como por el sistema.

El problema de los scripts es su variedad y que, en la mayoría de los casos, están compuestos por una o más estrategias de búsqueda. Por ejemplo, mientras que un usuario con una determinada necesidad informativa, hace *browsing* por la estructura de la base de datos para luego realizar una consulta – es decir, utilizaría dos estrategias – otro usuario, con la misma necesidad, podría realizarla directamente si cuenta con un mayor conocimiento de la materia – una sola estrategia –. Esto hace que sea necesario que el SRI tenga almacenadas la mayor cantidad posible de estrategias, de combinaciones entre ellas, y de scripts.

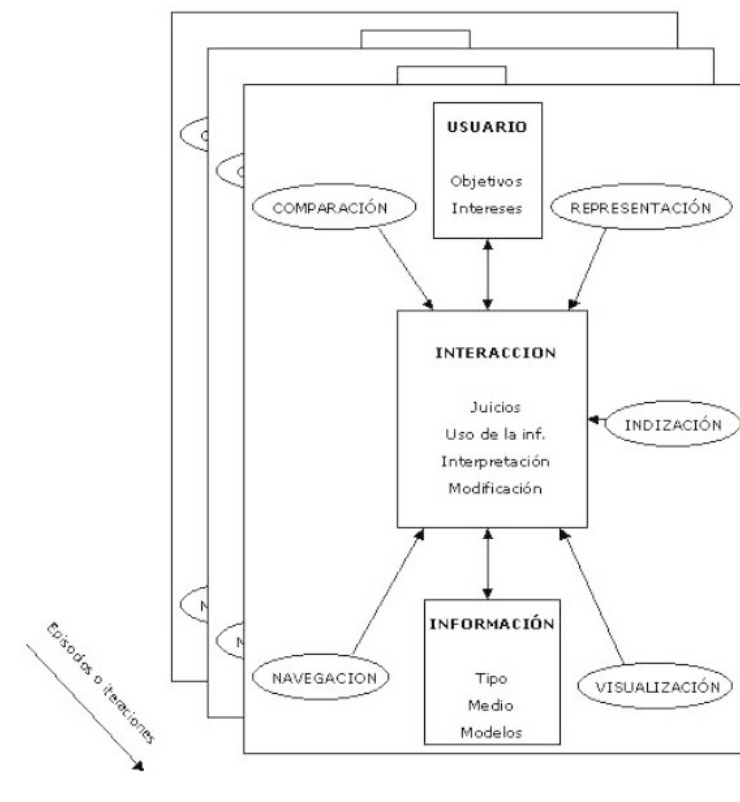
En resumen, cada consulta de usuario puede estar formada por uno o más episodios, cada episodio está formado por uno o más scripts, que a su vez están formados por una o más estrategias de búsqueda.

Las ventajas de este sistema son claras. Por un lado permite, de una forma evidente, la interacción entre el usuario y el SRI. Además, favorece la retroalimentación en los dos extremos, el usuario y el sistema. Y lo que es más importante, ha sido posible implementarlo en MERIT, una interfaz para una gran base de datos de información sobre publicaciones, proyectos y financiación europea[34], lo que implica la posibilidad de evaluarlo y de utilizarlo como base para comparar el rendimiento de las distintas bases de datos conectadas a él.

Los inconvenientes están ligados a los propios defectos del sistema, a saber, resulta muy difícil diseñar interfaces que permitan llevar a cabo, de una forma óptima,

FIGURA 3

Modelo Episódico de Recuperación de Información (Adaptado de Saracevic, 1996)



interacción y estrategias de búsqueda de información[2]. Por otra parte, y tal y como ya se ha indicado, el uso de los scripts limita, a la vez que condiciona en gran medida, el funcionamiento del sistema. Por último, el componente cognitivo del usuario está un poco abandonado ya que este modelo no tiene en cuenta aquellos aspectos contextuales, afectivos o sociales que condicionan la propia necesidad informativa y en los que *Ingwersen* tanto insiste[36].

MODELO ESTRATIFICADO

Para *Tefko Saracevic* la RI es uno más de los procesos que integran la interacción hombre-máquina (*Human-Computer Interaction* o HCI como lo llamaremos en adelante). Hasta tal punto considera este autor que HCI y RII son lo mismo que, para definir la RII, adopta la definición que propone *Storrs* para HCI[37]: (...) *el intercambio de información entre participantes, con el objetivo de cambiar o intercambiar el estado de conocimiento de uno mismo o del resto de participantes*. Pero va más allá y también adopta del HCI la definición de interacción: (...) *una interacción es un diálogo con el fin de cambiar el estado de*

conocimiento de uno o más participantes.

Saracevic llama modelo estratificado al sistema de RII propuesto[2] y considera que los sistemas de Ingwersen y Belkin, no contemplan de una forma adecuada el proceso de interacción, por lo que nunca podrían llegar a sustituir al modelo tradicional de RI.

Su sistema toma prestada la estructura de la *Teoría Estratificada de la Lingüística y la Comunicación*. Además, adopta y adapta muchas de las técnicas de los sistemas y de las investigaciones existentes sobre RII. Uno de sus objetivos principales es el de localizar e identificar los procesos de búsqueda de información de los usuarios, – a los que Belkin se refería como scripts – para incorporarlos al diseño de la interfaz, pues de ellos dependerá el éxito de la interacción entre sistema y usuario.

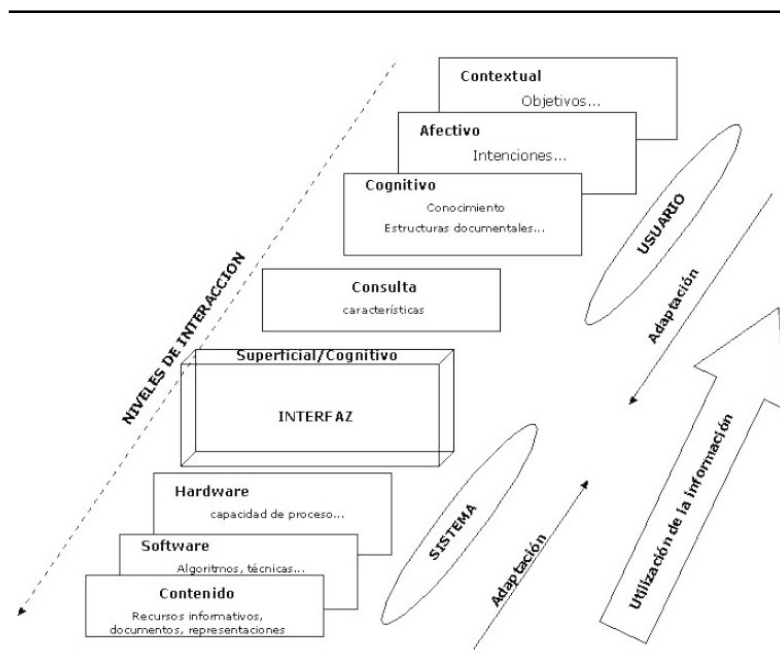
Como puede verse en la figura 4, los elementos o actores principales de esta propuesta son el sistema, la interfaz y el usuario. El primero está representado por el *hardware* y el *software* que constituyen el equipo informático y el SRI respectivamente, así como las representaciones documentales almacenadas en la base de datos. El segundo está compuesto por estructuras cognitivas y aspectos contextuales inherentes al propio usuario y al entorno en el que se mueve. La interfaz, que es la intermediaria entre los dos extremos, tiene como principal función la de favorecer la adaptación entre los distintos niveles y promover la interacción o diálogos – comunicación escrita, visual, gestual y cognitiva – , entre ambos.

Cada uno de los principales actores o elementos del sistema está dividido en estratos. Respecto del usuario, hay que considerar cuatro niveles: el superficial, el cognitivo, el afectivo y el contextual. Los estratos en torno a los que se estructura el sistema son: el superficial, el *hardware*, el *software* y el de contenido. En ambos casos, cada nivel implica diferentes elementos o procesos. Por último, la interfaz se divide en dos estratos, el superficial y el cognitivo.

La interacción se produce en el momento en que el usuario se pone en contacto con el SRI para utilizar la información y comienzan a dialogar por medio de la interfaz, todo ello con el único fin de influir y modificar el estado cognitivo del usuario. Cuando los estratos del

FIGURA 4

Modelo Estratificado de Recuperación de Información (Adaptado de Spink y Saracevic, 1998)



individuo entran en contacto, por medio de la interfaz, con los distintos estratos del sistema, no sólo se produce un proceso de equiparación total y/o parcial entre la consulta y las estructuras documentales almacenadas, tal y como ocurría en el modelo tradicional, sino que además tienen lugar otros fenómenos. Por un lado se producen diálogos entre el usuario y el sistema mediante la navegación o el *browsing* por la estructura de los documentos almacenados en el sistema. Además, se realizan interpretaciones y juicios cognitivos sobre los resultados de la consulta y el nivel de satisfacción de la necesidad informativa del usuario. Por otro lado, y en el caso de que la satisfacción no sea plena, se llevarán a cabo procesos de retroalimentación por relevancia para reorientar la consulta.

La interacción o diálogos anteriores prácticamente resolverían todos los problemas de la RII, si no fuese porque el individuo es un ser cambiante por naturaleza, porque la consulta puede estar mal definida y por todos aquellos aspectos sociales, contextuales, o emocionales, que acompañan al individuo y que hacen de la RII un proceso totalmente imprevisible.

Buscar la solución a estos problemas es algo que el HCI está intentando desde hace tiempo, y que Belkin pretendió resolver mediante una interfaz *adaptativa*. Esto es lo mismo que propone Saracevic, y para ello dota a la interfaz

de dos estratos. El primero, el superficial, tendría incorporadas todas aquellas características propias de los bibliotecarios o intermediarios de bases de datos, es decir, experiencia en el uso de la base de datos y en las técnicas de recuperación, trucos sobre su utilización, entre otros. El segundo estrato, el cognitivo, se utiliza fundamentalmente para que la interacción con el primer nivel sea efectiva, pareciéndose mucho al RMB de *Ingwersen*, o sea, un intermediario capaz de detectar la necesidad informativa del usuario, que le ayude a conceptualizarla, teniendo en cuenta las variables afectivas, sociales y contextuales que influyen en ella, y que le facilite el proceso de etiquetado o transcripción de la consulta a la interfaz.

El diseño de la interfaz, tal y como se propone en este modelo, hoy por hoy, aún no se ha conseguido. *Saracevic* y *Spink*, proponen para ello la observación de la interacción de los bibliotecarios o intermediarios de bases de datos con los usuarios para aprender de sus capacidades, destrezas y trucos, y extrapolarlas al diseño e implementación de la interfaz[3].

Las ventajas de este sistema se centran en su capacidad para poner de manifiesto de una forma clara que la RI es un proceso interactivo, así como en la localización e identificación de los elementos y procesos de la RII.

Por el contrario, sus inconvenientes son, en parte, los mismos que presentan los modelos estratificados en el campo de la lingüística y la comunicación, es decir, su descomposición no es fácil, y la descripción de los distintos niveles, así como sus relaciones, constituyen un aspecto muy criticado[2]. Además se denota una clara falta de integración entre sus elementos, dando la impresión de estar inacabado, esto puede deberse a que adopta y adapta elementos de otros modelos. Por otra parte, se presta mucha atención a los extremos del proceso interactivo, pero muy poca a la interfaz, que es donde precisamente este modelo descarga todo el proceso de interacción. Por último, los diálogos no aparecen claramente definidos ni en cuanto al tiempo ni en su duración, ni en cómo éstos pueden cambiar cuando lo hace la necesidad informativa y el estado cognitivo del usuario.

MODELO DE RETROALIMENTACIÓN INTERACTIVA

Una de las principales críticas que se les ha realizado a los sistemas interactivos, ha sido la de entender la interacción como algo inherente al sistema. Es decir, una vez diseñado el sistema interactivo y presentados sus

principales actores, no se especifica de una forma clara qué elementos ni procesos provocan y permiten la interacción entre los dos extremos del SRI.

Spink[5], mantiene que el principal responsable de la interacción, tanto en el modelo tradicional como en el interactivo, es la retroalimentación, pero teniendo muy en cuenta que la retroalimentación no es algo que tenga que estar ligado siempre a la relevancia, ni tiene porqué ser una función de la RI[38]. Para ella, la retroalimentación es mucho más, es el nombre general que se le da a una serie de elementos que permiten la interacción entre el sistema y el usuario.

Spink identificó diferentes tipos de retroalimentación a partir de un estudio realizado con 40 usuarios reales, con necesidades informativas reales, en donde se utilizó la base de datos *Dialog* y se contó con cuatro profesionales o intermediarios[5] para responder a las consultas de esos usuarios[3; 39]:

- *Retroalimentación por relevancia de contenido*: Se trata de la reformulación de una consulta teniendo en cuenta los juicios de relevancia que el usuario emite, como consecuencia de la respuesta que el sistema ofrece a una consulta anterior de éste. Este tipo de retroalimentación puede ser negativa o positiva, dependiendo de que según el juicio del usuario, los documentos sean relevantes o no. Representó el 40% de las interacciones entre el usuario/intermediario y el sistema.

- *Retroalimentación por relevancia de los términos*: Los usuarios reformulan la consulta incorporando nuevos términos. Éstos se obtienen de la observación y valoración de los resultados obtenidos en la consulta anterior. Sólo alcanzó el 8% de las interacciones entre el usuario/intermediario y el sistema.

- *Retroalimentación por magnitud de la respuesta*: Este tipo de retroalimentación la realiza el usuario en función del número de documentos que el sistema le ofrece como respuesta a su consulta. Puede ser tanto positiva, como negativa, dependiendo del exceso o defecto en el número de resultados. Representó el 45% de las interacciones entre el usuario/intermediario y el sistema. Fue la más elevada de todas las retroalimentaciones o interacciones.

- *Retroalimentación por revisión de consultas anteriores*: Muchas veces, a raíz de largas interacciones con el sistema, como consecuencia de una necesidad informativa poco clara o de un SRI poco eficiente, el usuario se encuentra perdido y no sabe qué estrategias de búsqueda ha utilizado.

Como solución mira en el histórico de su pantalla de búsqueda, en el caso de que la interfaz disponga de él. Sólo supuso el 6% de las interacciones entre el usuario/intermediario y el sistema.

– *Retroalimentación por revisión de términos utilizados en consultas anteriores*: Se trata del mismo caso que el anterior, con la diferencia de que aquí se buscan términos utilizados en consultas anteriores en lugar de estrategias. Su utilización en las interacciones es prácticamente despreciable, un 1%.

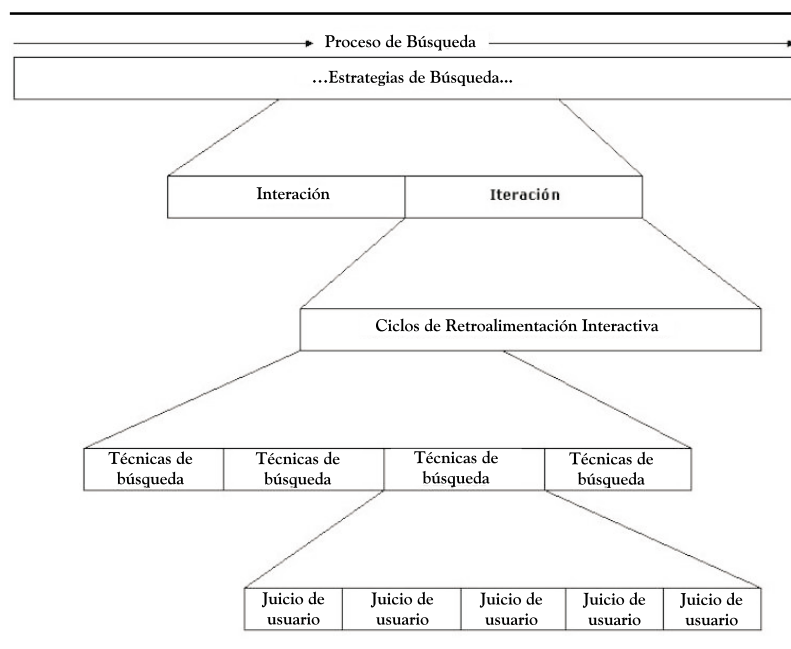
Una vez identificados y descritos los elementos sobre los que recae el peso de la retroalimentación, la autora del estudio, se basa en el sistema de RII de Saracevic contemplando sus distintos niveles de interacción. Para ello parte de la base de que el objetivo de toda interacción es el de cambiar el estado cognitivo de los intervinientes, e intenta explicar cómo se relacionan los distintos elementos que componen la RII mediante los distintos tipos de retroalimentación.

Como puede verse en la figura 5, cada estrategia de búsqueda puede estar formada por una o más iteraciones, es decir, una o más consultas al sistema con sus correspondientes respuestas. A su vez, cada iteración puede estar formada por uno o más ciclos de retroalimentación interactiva, donde un ciclo es (...) *una interacción que implica una consulta, un proceso (algorítmico) para obtener información del sistema, documentos (la respuesta), juicios por parte de un interprete (el usuario) y una acción consecuente (modificación de la consulta), que afecta a la salida documental o proceso*. De la misma forma, cada técnica de búsqueda, puede estar formada por distintas técnicas de búsqueda propias del usuario, destinadas a mejorar el proceso de recuperación, donde cada técnica de búsqueda espera una respuesta del sistema.

En definitiva, (...) *un proceso de búsqueda interactivo puede estar formado por una serie de estrategias de búsqueda, constituidas por una o más iteraciones y uno o más ciclos de retroalimentación interactiva de los tipos descritos antes* [39].

El sistema diseñado por Spink, no es tan amplio como pueden serlo el de Ingwersen, Belkin, o Saracevic. Es más, podría ser ampliamente criticado por no tener en cuenta aspectos tan importantes como la formación de la

FIGURA 5
Modelo de Retroalimentación Interactiva
(Adaptado de Spink, 1997)



necesidad informativa del usuario, su grado de incertidumbre o los elementos que influyen en ella y que pueden provocar el ASK del individuo, entre otros. Sin embargo, esta propuesta debe entenderse como la interfaz *adaptativa*, que se propone para el sistema estratificado.

Este sistema de retroalimentación interactiva presenta ventajas tales como la identificación de nuevos elementos que intervienen en el proceso de RII que, según Spink, son aquellos sobre los que recae todo el peso de la interacción. Por otro lado, estos nuevos elementos han sido descubiertos en un entorno real, con usuarios y necesidades informativas reales, lo que le da un gran valor desde el punto de vista cognitivo. Además se muestra la interacción como lo que realmente es, es decir, un proceso cíclico, en el que el estado de conocimiento del usuario cambia como consecuencia del uso de la información del sistema.

Los inconvenientes coinciden con los del sistema estratificado, pues se basa en éste. Un inconveniente añadido es el de incorporar la figura del intermediario en la parte experimental, pues aunque en definitiva tanto el usuario como el intermediario son usuarios finales, los resultados obtenidos no pueden ser extrapolados a los modelos ni SRI existentes, donde la figura del intermediario no existe. Como mucho, éste intermediario podría estar representado por la interfaz,

en el caso de que se le atribuyan todos los procesos y características anteriormente comentadas para los sistemas diseñados por *Inguersen* y *Belkin*. Pero esto no es posible, ya que tanto *Saracevic* como *Spink* consideran que, hoy por hoy, los únicos capaces de realizar los ciclos de retroalimentación interactiva son los humanos, y no las máquinas.

CONCLUSIONES

El modelo cognitivo, y los diferentes sistemas diseñados basados en él, surge como una alternativa al modelo tradicional al considerar que éste no representa de forma adecuada las necesidades informativas de los usuarios. Sin embargo, el modelo cognitivo se limita a realizar una aproximación casi exclusivamente teórica a las necesidades informativas de los usuarios desde diferentes perspectivas y, además, está exento de todo tipo de desarrollo e implementación práctica en entornos de usuarios. Prueba de ello, es que cualquiera de los diferentes diseños propuestos para los sistemas utilizan el, tan por ellos mismos criticado, modelo algorítmico para realizar los procesos de almacenamiento, selección y recuperación de información. Por otra parte, el modelo cognitivo, en su afán de estudiar las necesidades informativas, se centra excesivamente en el usuario, sin considerar que las conclusiones a las que llega corresponden a cada individuo concreto objeto de estudio, y que difícilmente son extrapolables a un grupo más amplio de individuos o incluso a una comunidad.

Es cierto que el modelo algorítmico o tradicional parece haber llegado a su límite de evolución. Las nuevas técnicas de recuperación desarrolladas a partir de él, desde el punto de vista estadístico, apenas suponen avances cuantificables con respecto a los resultados obtenidos mediante el uso de las técnicas desarrolladas durante los años 70. No obstante, este modelo sigue siendo el más utilizado en los actuales SRI comerciales, ya sea mediante técnicas de equiparación exacta o, como viene ocurriendo desde hace unos años, incorporando, técnicas de equiparación parcial.

Podría decirse que estamos en un tiempo de espera en el que el modelo tradicional y el cognitivo conviven, donde cada uno de ellos aporta lo mejor de sí mismo lo que ha provocado la aparición de un híbrido que puede denominarse *modelo interactivo de recuperación de información*. Este nuevo modelo intenta incorporar las virtudes de cada uno y eliminar sus inconvenientes, si bien hasta la fecha no se han logrado grandes éxitos dignos de mención. Sirvan como ejemplos los SRI descritos por

Belkin y sus colegas[18], donde se comprueba que los usuarios prefieren los sistemas basados en el modelo interactivo, pues requieren un menor esfuerzo cognitivo. No obstante, sus resultados, evaluados por los propios usuarios mediante técnicas algorítmicas, ofrecen unos valores sensiblemente inferiores a los obtenidos con los SRI basados en el modelo tradicional. Además, una evaluación realista de los SRI interactivos debe ser multidimensional ya que estos modelos han creado la necesidad de nuevas medidas de la eficacia de la recuperación, aún no consensuadas, que contemplen todas las facetas de la interacción con el usuario. El uso de estas medidas origina un interesante debate en torno a los propios conceptos, a su interpretación y a la aplicabilidad de las mismas[40].

En este sentido, *Kuhn*[41], señala como una de las condiciones indispensables para la existencia de un paradigma científico, los ejemplares compartidos (o *shared exemplars*), es decir, modelos de experimentación especialmente exitosos que permiten detectar empíricamente un paradigma y dotan de contenido y fuerza explicativa a una teoría científica, además de establecer los métodos para investigar un determinado problema. Mientras que en el modelo tradicional, los tests de Cranfield, el Smart de Salton o la Conferencia TREC cumplen esta función, se hace extremadamente difícil señalar una prueba o un estudio particulares como modelo de la aproximación cognitiva.[42]

Quizá sea el momento de proponer enfoques alternativos que complementen de manera más eficaz a los actuales modelos de RI. Enfoques que no tienen por qué desplazar ni ser excluyentes con los diferentes desarrollos aquí comentados. Nuestra propuesta, se basa en el acceso y en la RI a partir de la simplificación de las representaciones documentales mediante técnicas de visualización, tal y como proponen *Small*[43], *White* y *McCain*[44],[42] o *Ding*[45], entre otros, acompañada de un análisis del dominio, o área de conocimiento, desde un punto de vista social, funcionalista y real, al estilo de las propuestas de *Hjorland* y *Albrechtsen*[46].

REFERÊNCIAS

1. CLEVERDON, C. The Cranfield tests on index languages devices. In: SPARCK, Jones K.; WILLET, P. (Ed.). *Readings in information retrieval*. San Francisco : Morgan Kaufmann, 1997.
2. SARACEVIC, T. Modeling interaction in information retrieval (IR): a review and proposal. In: ANNUAL MEETING OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE, 1996. **Proceedings...** [S. l.], ASIS, 1996. p. 3-9.
3. SPINK, A.; SARACEVIC, T. Interaction in information retrieval: selection and effectiveness of search terms. *Journal of the American Society for Information Science*, v. 48, n. 8, p. 741-761, 1997.
4. MOYA ANEGÓN, F. Los sistemas integrados de gestión bibliotecaria. Madrid : Anabad, 1995.
5. SPINK, A. Study of interactive feedback during mediated information retrieval. *Journal of the American Society for Information Science*, v. 48, n. 5, p. 382-394, 1997.
6. SALTON, G.; MCGILL, M. J. The smart and sire experimental retrieval systems . In: _____; _____. Introduction to modern information retrieval. New York : McGraw-Hill, 1983. p. 118-151.
7. _____; _____. Introduction to modern information retrieval. New York : McGraw-Hill, 1983.
8. HARMAN, D. Relevance feedback revisited. In: ANNUAL INTERNATIONAL CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL, 15., 1992, Copenhagen. **Proceedings...** Copenhagen : ACM, 1992. p. 1-10.
9. SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. *Journal of the American Society for Information Science*, v. 26, n. 11-12, p. 321-343, 1975.
10. MIZZARO, S. Relevance: the whole history. *Journal of the American Society for Information Science*, v. 48, n. 9, p. 810-832, 1997.
11. BELKIN, N. J. The cognitive viewpoint in information science. *Journal of Information Science*, n. 16, p. 11-15, 1990.
12. _____; ODDY, R. N.; BROOKS, H. M. Ask for information retrieval: Part I: Background and theory. *Journal of Documentation*, v. 38, n. 2, p. 61-71, 1982.
13. _____; _____; _____. Ask for information retrieval: Part II. Results of a design study. *Journal of Documentation*, v. 38, n. 3, p. 145-164, 1982.
14. BATES, M. J. The design of browsing and berrypicking techniques for the online search interface. *Online Review*, v. 13, n. 5, p. 407-424, 1989.
15. ELLIS, D. A behavioural approach to information system design. *Journal of Documentation*, v. 45, n. 3, p. 171-212, 1989.
16. EFTHIMIADIS, E. N. Interactive query expansion: a user-bases evaluation in a relevance feedback environment. *Journal of the American Society for Information Science*, v. 51, n. 11, p. 989-1003, 2000.
17. ROBERTSON, S. E.; WALKER, S. ; HANCOCK-BEAULIEU, M. M. Large test collection experiments on an operational, interactive system: Okapi and Trec. *Information Processing and Management*, v. 31, p. 345-360, 1995.
18. BELKIN, N. J. *et al.* Iterative exploration, design and evaluation of support for query reformulation interactive information retrieval. *Information Processing and Management*, v. 37, p. 403-434, 2001.
19. INGWERSEN, P. Information retrieval interaction. London : Taylor Graham, 1992.
20. KUHLETHAU, C. C. A principle of uncertainty for information seeking. *Journal of Documentation*, v. 49, n. 4, p. 339-355, 1993.
21. INGWERSEN, P. Polyrepresentation of information needs and semantic entities: elements of a cognitive theory for information retrieval interaction. In: ANNUAL INTERNATIONAL CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL, 17., 1994, Berlin. **Proceedings...** Berlin : Springer Verlag, 1994, p. 101-110.
22. TANG, R.; SOLOMON, P. Toward an understanding of the dynamics of relevance judgment: an analysis of one person's search behavior. *Information Processing and Management*, v. 34, n. 2/3, p. 237-256, 1998.
23. INGWERSEN, P. Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of Documentation*, v. 52, n. 1, p. 3-50, 1996.
24. SUTCLIFFE, A. E. M. Towards a cognitive theory of information retrieval. *Interacting with Computers*, v. 10, p. 321-351, 1998.
25. SPINK, A.; LOSEE, R. M. Feedback in information retrieval. *Annual Review of Information Systems and Technology*, n. 31, p. 33-77, 1996.
26. SALTON, G.; BUCLEY, C. Improving retrieval performance by relevance feedback. *Journal of the American Society for Information Science*, v. 41, n. 4, p. 288-297, 1990.
27. INGWERSEN, P.; WILLET, P. An introduction to algorithmic and cognitive approaches for information retrieval. *Libri*, n. 45, p. 160-177, 1995.
28. ROBERTSON, S. E. The probability ranking principle in IR. *Journal of Documentation*, v. 33, n. 4, p. 294-304, 1977.
29. BELKIN, N. J. *et al.* The effect of multiple query representations on information retrieval system performance. In: ANNUAL INTERNATIONAL CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL, 16., 1993, New York. **Proceedings...** New York : ACM, 1993. p. 339-346.
30. HAINES, D.; CROFT, W. B. Relevance feedback and inference networks. In: ANNUAL INTERNATIONAL CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL, 16., 1993. **Proceedings...** New York : ACM, 1993. p. 2-11.
31. LEE, J. H. Combining multiple evidence from different relevance feedback methods. *Information Processing and Management*, v. 34, n. 6, p. 681-691, 1998.
32. RUTHVEN, I. Incorporating aspects of information use into relevance feedback. 1999. Disponível em: < <http://ir.dcs.gla.ac.uk/papers/Pdf/ruthven2000a.pdf>>. Acesso em: 1 jul. 2001.
33. _____; LAMAS, M.; RIJSBERGEN, C. J. Van. Retrieval through explanation: an abductive inference approach to relevance feedback. 1999. Disponível em: <<http://ir.dcs.gla.ac.uk/papers/pdf/ruthven99a.pdf>> Acesso em: 1 jul. 2001.
34. BELKIN, N. J. *et al.* Cases, scripts, and information-seeking strategies: on the design of interactive information retrieval systems. *Expert Systems with Applications*, n. 9, p. 379-395, 1995.
35. _____; CROFT, W. B. Information filtering and information retrieval: two side of the same coin? *Communications of the ACM*, v. 35, n. 12, p. 29-38, 1992.
36. ROBINS, D. Interactive information retrieval: contexts and basic notion. *Informing Science*, v. 3, n. 2, p. 57-62, 2000.
37. STORRS, G. A conceptualization of multiparty interaction. *Interacting with Computers*, v. 6, n. 2, p. 173-189, 1994.

38. SPINK, A.; GREISDORF, H.; BATEMAN, J. From highly relevant to not relevant: examining different regions of relevance. *Information Processing and Management*, v. 34, n. 5, p. 599-621, 1998.
39. _____; Saracevic, T. Human-computer interaction in information retrieval: nature and manifestations of feedback. *Interacting with computers*, v. 10, p. 249-267, 1998.
40. OLVERA, M. D. Evaluación de sistemas de recuperación de información: aproximaciones y nuevas tendencias. *El Profesional de la Información*, v. 8, n. 11, p. 4-14, 1999.
41. KUHN, T. S. La estructura de las revoluciones científicas. México D.F.: Fondo de la Cultura Económica, 1990.
42. WHITE, H. D.; MCCAIN, K. W. Visualization of literatures. *Annual Review of Information Systems and Technology*, v. 32, p. 99-168, 1997.
43. SMALL, H. Visualizing science by citation mapping. *Journal of the American Society for Information Science*, v. 50, n. 9, p. 799-813, 1999.
44. WHITE, H. D.; MCCAIN, K. W. Visualizing a discipline: an author co-citation analysis of information science, 1972-1995. *Journal of the American Society for Information Science*, v. 49, n. 4, p. 327-355, 1998.
45. DING, Y.; CHOWDHURY, G. G.; FOO, S. Bibliometric cartography of information retrieval research by using co-word analysis. *Information Processing & Management*, v. 37, p. 801-817, 2001.
46. HJORLAND, B. Birger; ALBRECHTSEN, Hanne. Toward a new horizon in information science: domain analysis. *Journal of the American Society for Information Science*, v. 46, n. 6, p. 400-425, 1995.